



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Self-Excitation in the Solar Flare Waiting Time Distribution

Citation for published version:

Ross, GJ 2020, 'Self-Excitation in the Solar Flare Waiting Time Distribution', *Physica a-Statistical mechanics and its applications*, vol. 556, 124775. <https://doi.org/10.1016/j.physa.2020.124775>

Digital Object Identifier (DOI):

[10.1016/j.physa.2020.124775](https://doi.org/10.1016/j.physa.2020.124775)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Physica a-Statistical mechanics and its applications

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Self-Excitation in the Solar Flare Waiting Time Distribution

Gordon J Ross

*Department of Mathematics, University of Edinburgh,
Edinburgh EH9 3FD, UK*

*

(Dated: June 1, 2020)

Solar flares release high amounts of energy into the solar system and can negatively impact earth based systems through their effects on satellites and power systems. It is hence important to understand and forecast their occurrence. The solar flare waiting time distribution (WTD) defines the amount of time which elapses between the occurrence of successive flares and hence provides a starting point for forecasts and risk assessment. Previous research has hypothesised that the observed WTD can be derived from a simple model which posits that flares follow a nonstationary Poisson process. This Poissonian assumption has implications for fundamental physical theories about the origin of flares, since it is a direct consequence of the widely studied avalanche model. However in this paper we call the Poissonian assumption into question, by showing that the occurrence of solar flares seems to have a substantial amount of burstiness and self-excitation that continues to exist even when controlling for the solar cycle. This leads to a strong non-Poissonian dependence between the occurrence time of successive flares.

I. INTRODUCTION

Solar flares are high energy bursts of radiation that are frequently emitted from the sun, and are considered to be one of the most significant types of space weather which can adversely affect important systems such as satellites and radio transmissions [1, 2]. As such, it is important to understand the physical and statistical processes which underlie their occurrences. Much interest has focused on modelling the waiting-time distribution (WTD) of flares, which is defined as the distribution of the time that elapses between successive flares, i.e. $p(\Delta_i)$ where $\Delta_i = t_{i+1} - t_i$ and t_i is the occurrence time of the i^{th} flare.

It is generally believed that the flare WTD seems to follow a power-law distribution [3, 4], although there is disagreement as to the underlying cause. One potential explanation is that the power-law arises due to the self-similar nature of solar flares [5–7]. However in an important series of papers, [3, 8, 9] argued that the power-law distribution could also be due to flare occurrences following a nonstationary Poisson process with the process rate undergoing frequent change. If true, then this is a significant finding since the resulting Poissonian statistics result in a strong independence between the occurrence times of successive flares with no residual clustering, and support a non-stationary avalanche model for their ultimate origin [3].

As well as its scientific interest, an improved understanding of solar flare occurrence times also has practical implications. Solar flares and other extreme forms of space weather can have a severe impact on electrical devices such as radios and GPS satellites [1] [2]. As such, building more sophisticated occurrence models can po-

tentially improve the quality of predictions and forecasts, as well as leading to improved risk quantification [10–12].

However, despite much subsequent work that has examined the non-stationary Poisson hypothesis [13, 14], there has only been limited attention paid to whether this is a truly adequate model for solar flare occurrences. In particular, the question of whether alternative models may have a higher degree of empirical support has not received much attention, although there are exceptions [5]. In this paper, we argue that the high degree of clustering in solar flare occurrence times means that they are better modelled using a self-exciting point process which allows for both seasonality in the background rate of occurrences due to the 11-year solar cycle, while also showing residual burstiness with flares occurring in clusters. The self-exciting process we use is a type of Hawkes process, which has previously been shown to be a good model for many other empirical phenomena such as tsunamis [15], hurricane landfalls [16] and even human-made processes such as financial markets [17] [18] and social networks [19]. Interestingly, the Hawkes process is also widely used for modelling the occurrence of earthquakes, where it is known as the ETAS model [20, 21]. The fact that both earthquakes and solar flares are well-described by the same statistical process builds on previous universality results between these two seemingly separate type of events, which was previously noted by [22]. Our results hence call into question the assumption that even a highly non-stationary Poisson process is capable of capturing the intricacies of solar flare occurrences.

The rest of the paper is arranged as follows: in Section II we describe the solar flare catalog that will be analyzed, and in Section III we examine whether Poisson process models can reproduce the empirically observed waiting time distribution. Section IV explores the residual burstiness/self-excitation which is present solar flare occurrences which does not seem to be captured by Pois-

* gordon.ross@ed.ac.uk

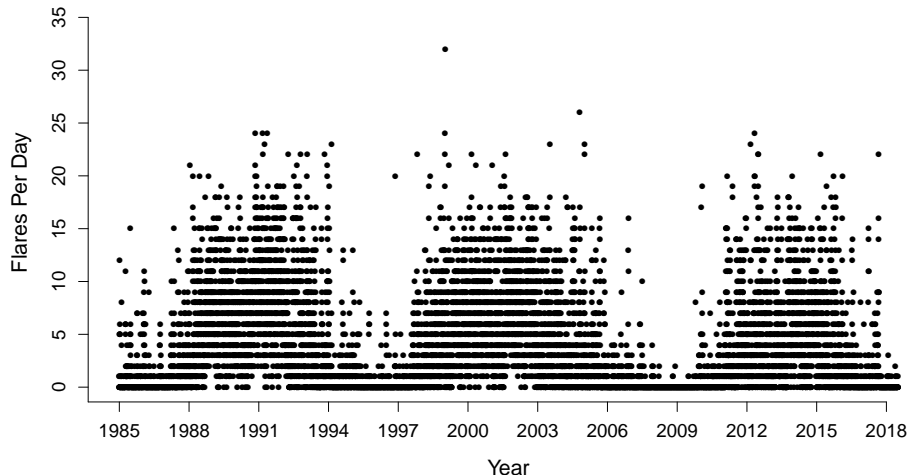


FIG. 1: Number of daily solar flares of class C or higher in the GOES catalog, for each day during the period 1985-2018.

son process models. Finally conclusions are drawn in Section VI.

II. DATA

We analyze the soft X-ray flare catalog that collected by the Geostationary Orbital Environmental Satellites (GOES) between the years 1985 and 2018. This catalog was previously used by [3], [5], [4] and others in their respective studies of solar flare waiting time distributions, and is available from http://hec.helio-vo.eu/hec/hec_gui.php. We follow the procedure of [3] and only consider flares that are class C1.0 or higher (i.e. class C, M or X) to reduce the amount of missing data that is caused by low magnitude flares being undercounted during periods of solar maximum. This leaves a total of $n = 39,754$ flares in the catalog. Figure 1 plots the number of flares that occurred each day during the 34 year observation window. It can be seen that flare activity is strongly periodic, with a period of roughly 11 years which tracks the solar cycle [23]. Note that this discretization into the number of flares per day is only used here for visualization; we will work in continuous time when fitting models to the data.

Let t_1, \dots, t_n be the occurrence time of each flare, measured in days. Figure 2 shows the empirical density (continuous histogram) of the logarithm of the waiting times i.e. $\log(t_{i+1} - t_i)$, for the catalog. The best fitting Normal distribution is superimposed on top, and it can be seen that it gives a relatively good, albeit imperfect, fit. This implies that the waiting times Δ_i have a distribution which is close to lognormal, although distinguishing between the lognormal and power-law distributions can

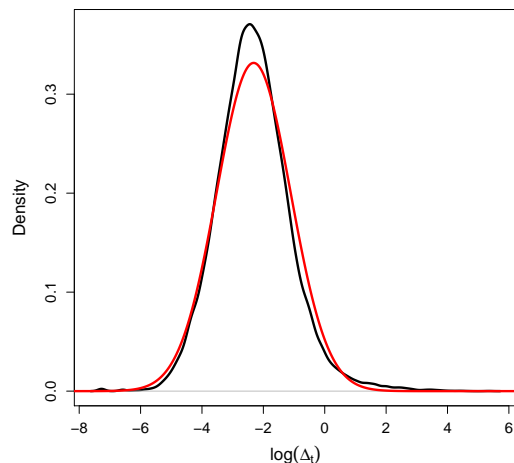


FIG. 2: Empirical density of the (log) solar flare waiting times with the best fitting Normal distribution superimposed in red. An approximate Normal fit to the log density implies that the waiting time distribution is close to lognormal.

be a difficult task [24]. Fitting a power law to the data using the standard maximum likelihood approach from [24] yields a power law exponent of $\alpha = -2.21$ for waiting times above 1.2 days. A nonparametric bootstrap [25] can be used to construct a 95% confidence interval for this, yielding the interval $\alpha \in (2.16, 2.23)$.

III. POISSON PROCESS

Suppose that n events occur in the time interval $[0, T]$, at times t_1, \dots, t_n . These events can be modelled as a point process with intensity function $\lambda(t)$ which defines the (limiting) instantaneous probability of an event occurring at each point in time. The simplest type of point process is a homogenous Poisson process (HPP) where $\lambda(t) = \lambda$ is constant over time. In this case, the number of events on each interval of length δ follows a Poisson($\delta\lambda$) distribution:

$$p(k \text{ events in } [t, t + \delta]) = \frac{(\delta\lambda)^k e^{-\delta\lambda}}{k!} \quad (1)$$

and the waiting times follow an Exponential(λ) distribution [26]. The main assumption of the HPP is that events occur uniformly over time, i.e. that each interval of length δ will have the same expected number of events as all other intervals of equal length, and that the events in each interval occur independently. Such a strong assumption of homogeneity is unrealistic for solar flares, since Figure 1 shows considerable variation in the occurrence rate over time. To take this into account, [3] and [8] proposed modelling $\lambda(t)$ as a time-varying step function, which involves dividing $[0, T]$ into a union of disjoint intervals, with $\lambda(t)$ constant on each interval. Suppose there are m such intervals and let $\tau_1, \tau_2, \dots, \tau_{m-1}$ denote the (unknown) points at which $\lambda(t)$ changes value. This leads to the specification:

$$\lambda(t) = \begin{cases} \lambda_1 & \text{if } 0 \leq t < \tau_1 \\ \lambda_2 & \text{if } \tau_1 \leq t < \tau_2 \\ \dots & \\ \lambda_m & \text{if } \tau_{m-1} < t \end{cases} \quad (2)$$

where $\lambda_1, \dots, \lambda_{m+1}$ are constants which define the value of $\lambda(t)$ on each interval. This model was shown in [3] to give a good qualitative fit to the observed waiting time distribution, although some discrepancies were observed. We will refer to this approach as the change point Poisson process (CPPP).

Given the previous Figure 1, it is questionable whether modelling $\lambda(t)$ as a step function is the most parsimonious approach. Solar flare activity seems to follow a sinusoidal pattern, as would be expected given the quasi-periodic nature of the 11 year solar cycle. As such, we instead propose modelling $\lambda(t)$ as a seasonal/periodic function with a periodicity of 11 years:, which we will call the seasonal Poisson process (SPP). Under this specification, $\lambda(t)$ will be defined as:

$$\lambda(t) = A + B \sin(\omega t) + C \cos(\omega t), \quad \omega = \frac{2\pi}{D} \quad (3)$$

where D is the length of the average solar cycle over the observed period, which we expect to be roughly equal to

11 years [23]. However since the solar cycle may not be exactly equal to 11 years, we treat D as a free parameter to be estimated along with the unknown A, B, C .

We hence have three possible models to describe solar flare occurrence, corresponding to constant intensity, the change point step function model, and the seasonal model. To choose the best model we will use maximum likelihood estimation penalized by the Bayesian Information Criterion, which is one of the most widely used approaches for comparing probabilistic models [27–29]. For a given intensity function $\lambda(t)$ and observed data t_1, \dots, t_n , the log-likelihood of a general point process is given by [26]:

$$LL(t_1, \dots, t_n | \lambda(t)) = \sum_{i=1}^n \log \lambda(t_i) - \int_0^T \lambda(s) ds \quad (4)$$

If we simply choose the model which has the highest maximum likelihood, then models with a large number of free parameters will be unfairly advantaged which results in over-fitting. As such it is common to penalize the maximum likelihood based on the number of free parameters in the model. The most widely used penalty is the BIC which is derived from an asymptotic approximation to the maximum likelihood and is given by:

$$BIC(t_1, \dots, t_n | \lambda(t)) = LL(t_1, \dots, t_n | \lambda(t)) - 0.5k \log(n) \quad (5)$$

where k is the number of parameters in $\lambda(t)$. For the homogenous Poisson process $k = 1$ since λ is the only parameter. The seasonal process has 4 free parameters (A, B, C, D), while the number of free parameters in the change point specification is equal to $2m + 1$.

Fitting the three models to the GOES solar flare data is done by numerically maximising the likelihood/BIC given the observed flares. For the homogenous and seasonal models, this can be done using direct maximization. The change point specification is more difficult since the number of change points $m - 1$ is also unknown and must be estimated at the same time. Since there are 2^n possible change point configurations, this cannot be done by brute force optimization. Instead, we use the PELT (Pruned Exact Linear Time) algorithm from [30] which utilises dynamic programming to convert the optimization problem into an $O(n^2)$ search operation which is a computationally easy task for the solar flare data.

Figure 3 plots the solar flare catalog with the best fitting change point and seasonal specifications for $\lambda(t)$ superimposed, which were found by maximising the BIC for each model. Unsurprisingly, the change point approach requires a large number of segments in order to give a good fit to the catalog. When the change point model was fit to the data using the above approach, 376 different segments were found, with each containing an average of 106 flares. The large number of free parameters needed by the change point model reflects the fact that the data has a strong sinusoidal aspect and is not

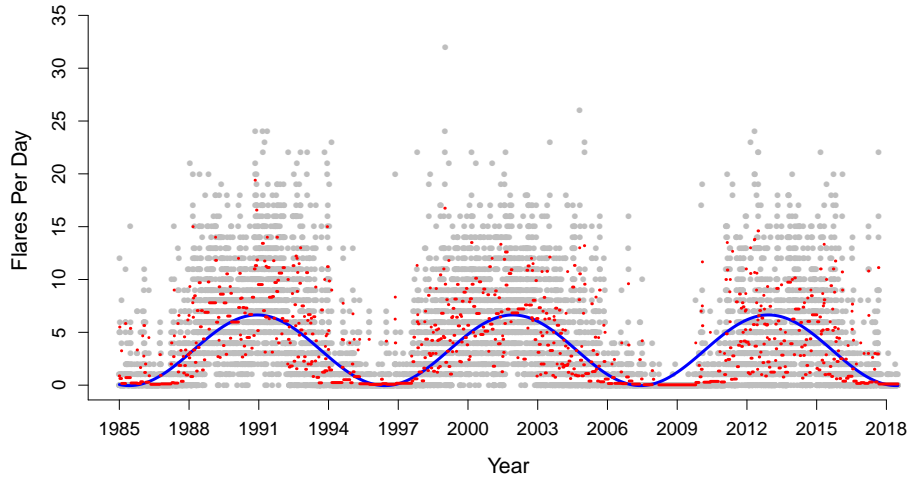


FIG. 3: Number of daily solar flares of class C or higher in the GOES catalog, for each day during the period 1985-2018. The best fitting change point Poisson process is superimposed in red, and the best fitting seasonal specification $\lambda(t) = A + B \sin(2\pi/11 \times 365) + C \cos(2\pi/11 \times 365)$ is superimposed in blue.

with a period of 4140 days (11.34 years).

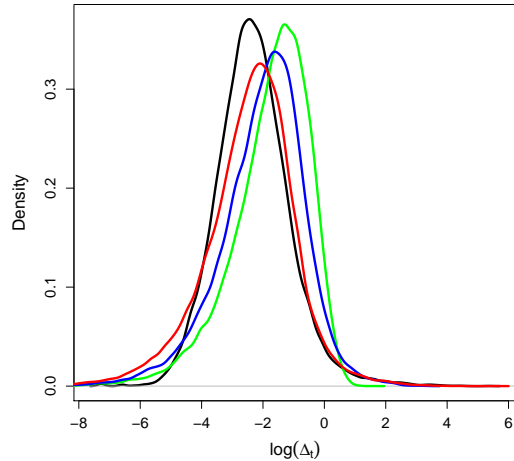


FIG. 4: The black line shows the empirical waiting time distribution $p(\log(\Delta))$ from the solar flare catalog. The colored lines show the waiting time distributions that correspond to the fitted homogenous Poisson process (green), change point Poisson process (red), and seasonal Poisson process (blue).

easily captured by a step function. In contrast, the seasonal approach gives a good fit to the catalog despite only requiring 4 parameters rather than 376. The best fitting seasonal specification was:

$$\lambda(t) = 3.30 - 0.8 \sin(\omega t) - 3.00 \cos(\omega t) \quad (6)$$

Table I shows the resulting likelihoods and BICs for each model. It can be seen that despite the very large number of additional free parameters, the change point specification fits the data substantially better than the seasonal specification even after penalizing for model complexity through the BIC. This suggests that solar flare occurrences have more complex local structure than can be captured by a simple seasonal specification. To investigate further, Figure 4 assesses whether any of the Poisson process models are able to reproduce the empirically observed waiting time distribution. This is done by plotting the unconditional distribution $p(\Delta)$ under each of the three fitted models. Unsurprisingly, the homogeneous Poisson process completely fails to reproduce the observed waiting time distribution. The seasonal specification is a substantial improvement, however the best match is given by the change point approach. This validates the observation from [3] that the change point approach seems to capture qualitative features of observed flare catalogs. However the fact that so many free parameters (376) are needed feels unsatisfactory and suggests a severe lack of parsimony. In the next section we will show that a small modification to the seasonal specification produces a model which seems to capture the observed solar flare occurrences to an even greater degree than the change point model, despite having substantially fewer parameters.

	BIC
Homogenous PP	7113
Seasonal PP	19621
Change-point PP	32115

TABLE I: BIC values for each Poisson process specification. Higher numbers indicate a better fit to the data.

IV. SELF-EXCITATION

Regardless of which of the above specification are used for $\lambda(t)$, the basic feature of the Poisson process is that events are indepently distributed according to $\lambda(t)$. This implies that there is no long-range dependence, burstiness, or clustering in the event times. However it is known that dependent, bursty behavior is common in many empirically observed processes such as earthquakes [20] [21], tsunamies [15], hurricane landfalls [16] and even human-made processes such as financial markets [17] [18] and social networks [19]

The Hawkes process [31] is a self-exciting point process which has been used to model bursty ih many of the above physical processes, and hence has some degree of universality. When used to model earthquakes, it is commonly known as the ETAS model. For a general Hawkes process, the intensity function is given by:

$$\lambda_{ETAS}(t) = \lambda_0(t) + \sum_{t_i < t} \kappa g(t - t_i) \quad (7)$$

The function $\lambda_0(\cdot) > 0$ is known as the background process and specifies the baseline intensity, and the function $g(\cdot) > 0$ is known as the triggering kernel and is a probability distribution that integrates to 1. The intuition behind the Hawkes process is that the $\lambda_0(t)$ function describes the ‘average’ behavior of the process, including any seasonal/periodic aspects. Similar to the previous section, it can be chosen as $\lambda_0(t) = \lambda$ to give a constant baseline intensity, or as $\lambda_0(t) = A + B \sin(\omega t) + C \cos(\omega t)$ to incorporate seasonality.

Next, the triggering kernel and scaling constant $\kappa > 0$ determine the contribution of each event previous t_i to the intensity at time t . This means that the total intensity at each time t is a sum of the baseline intensity, and an extra contribution that comes each previous event. Typically the $g(\cdot)$ function is chosen to be be monotonic decreasing and greater than zero everywhere. This implies that each event causes the process intensity to temporarily increase, producing local clusters of events but that this effect fades out over time. The choice of $g(\cdot)$ is application dependent and several parametric forms have been proposed in the literature, with popular choices including exponential decay:

$$g(z) = \beta \exp(-\beta z) \quad (8)$$

which corresponds to a short memory process where the effect of each event quickly fades out, or Lognormal decay:

$$g(z) = \frac{1}{z\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log(z) - \mu)^2}{2\sigma^2}\right) \quad (9)$$

where the additional parameter allows for a more flexible specification that can incorporate longer term memory.

The Hawkes process can equivalently be represented as a branching process, as first noted by [31]. At each time t , suppose n_j events occurred prior to t . From Equation 7, the conditional intensity at time t is a linear superposition of $n_j + 1$ independent nonhomogenous Poisson processes, where the first is the background process contributing intensity $\lambda_0(t)$ and the remainder are indexed by each of the n_j previous events, each contributing intensity $g(t - t_j)$ respectively. Since these processes are independent, each event time t_i can be assumed to have been generated either by the background process $\lambda_0(\cdot)$ or by the $\kappa g(\cdot)$ processes triggered by one of the previous events.

To aid interpretation of the branching process representation, consider the following algorithm for simulating a sequence of events from the Hawkes process on $[0, T]$: first simulate a collection of events from an nonhomogenous Poisson process on $[0, T]$ with intensity function $\lambda_0(t)$. Suppose n^0 events are generated and denote these by $t_1^0, t_2^0, \dots, t_{n^0}^0$. These events are called the baseline events. Next for each baseline event t_i^0 , simulate events from a further nonhomogenous Poisson process on $[t_i^0, T]$ with intensity function $\kappa g(t - t_i^0)$. Suppose n^1 new events are generated in total and denote these by $t_1^1, t_2^1, \dots, t_{n^1}^1$. For each event t_i^1 , simulate events from a further nonhomogenous Poisson process on $[t_i^1, T]$ with intensity function $g(t - t_i^1)$. Repeat this procedure with each generated event triggering a further nonhomogenous Poisson process until no more events are generated (which will eventually happen with probability 1 since $\int_0^T g(t) dt < \infty$). Finally, order all the $n^0 + n^1 + \dots = n$ events generated by this procedure as t_1, t_2, \dots, t_n . The collection of generated events is then a sample from the Hawkes process on $[0, T]$ with conditional intensity given in Equation 7.

V. RESULTS

We postulate the following Hawkes process as a potential generating mechanism for solar flare data:

$$\lambda(t) = \lambda_0(A + B \sin(\omega t) + C \cos(\omega t)) + \sum_{t_i < t} \kappa LN(t - t_i) \quad (10)$$

where LN denotes the lognormal distribution from Equation 9. The Lognormal choice for the trigger kernel is motivated by Figure 2 which previously showed an approximate Normal fit to the logarithm of the waiting

$A = 3.20 \pm 0.016,$	$B = 0.064 \pm 0.002,$	$C = -3.08 \pm 0.017$
$\lambda_0 = 0.11 \pm 0.004,$	$\kappa = 0.89 \pm 0.006,$	$\beta = 1.22 \pm 0.023$
$\lambda_0 = 0.10 \pm 0.005,$	$\kappa = 0.90 \pm 0.007,$	$\mu = -0.73 \pm 0.023,$
		$\sigma = 1.22 \pm 0.017$

TABLE II: Parameter estimates and standard errors for the seasonal model, and Exponential/Lognormal Hawkes processes.

times. This results in a model with only 7 free parameters: the A, B, C parameters which characterise the seasonal baseline intensity, the seasonal period D , and the 3 free parameters κ, μ, σ corresponding to the self-exciting element of the Lognormal Hawkes process. Note that λ_0 is not a free parameter since this model could be reparameterised to remove it, by multiplying its value into A, B and C . We have chosen to parameterise the model as above so that the same values of (A, B, C) can be used as in the seasonal model, to make the resulting comparison more interpretable. Again, this model was fit to the data by maximising the process likelihood function from Equation 4. Table III shows the resulting fit for all models, with those from the previous section also included for comparison. Note that we also considered a Hawkes specification using an Exponential kernel (Equation 8) for completeness.

	BIC
Homogenous PP	7113
Seasonal PP	19622
Change-point PP	32115
Exponential Hawkes	31928
Lognormal Hawkes	32476

TABLE III: BIC values for each fitted model. Higher numbers indicate a better fit to the data.

From the results, we can see that lognormal Hawkes process provides an extremely good match to the solar flare data. Compared to the seasonal Poisson process which includes the sin/cos terms but no self-excitation, the Hawkes process gives a drastically better fit. This suggests that there is substantial local structure in solar flare waiting times which calls the Poissonian assumption into dispute. The Hawkes process also gives a better fit to the data than the change point approach despite only having 7 free parameters compared to 376. This shows that it is capturing a substantial degree of structure which is being missed by even such a rapidly changing step function, suggesting that self-excitation may be a genuine non-Poissonian aspect of solar flare occurrences. We hence interpret these findings as casting serious doubt on the claim that solar flares can be described by structureless Poisson processes, even with the addition of substantial non-stationary components through the piecewise specification of $\lambda(t)$, which contradicts previous claims by (e.g.) [3] that flares can be

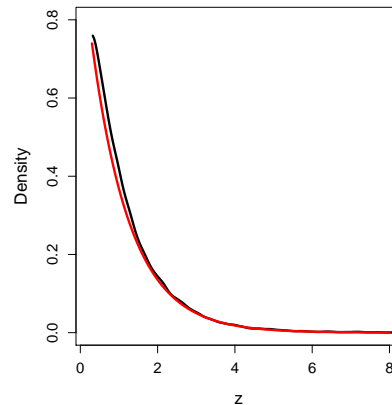


FIG. 5: The black line shows the distribution of the time-rescaled residuals $z_i - z_{i-1}$ for the Lognormal Hawkes process, which will be Exponential(1) if the model is correct. The red line shows the true Exponential(1) distribution.

viewed as independent random events.

For reference, Table II shows the parameter estimates and standard errors for the Hawkes process models, where the standard errors were computed based on the inverse Hessian matrix evaluated at the maximum likelihood estimate. Additionally, it is interesting to ask whether the Lognormal Hawkes process gives an objectively good fit to the solar flare data, rather than just giving a relatively better fit than the other models which were considered. For this purpose, we can perform a standard goodness-of-fit test for point processes as described in (e.g.) [32]. For each observation t_i , define $z_i = \int_0^{t_i} \hat{\lambda}(s) ds$ where $\hat{\lambda}(\cdot)$ is the maximum likelihood estimate of the point process intensity function. These z_i are known as the time-rescaled residuals, and it can be shown that they are realisations of a homogenous Poisson process with unit intensity [32]. As such, the differences $z_i - z_{i-1}$ are independent and identically distributed with distribution Exponential(1). Figure 5 plots the distribution of these z_i variables for the Lognormal Hawkes model, with the distribution of the Exponential(1) superimposed. It can be seen that these are extremely close, suggesting that the simple Lognormal model is capturing most of the factors which drive solar flare occurrences.

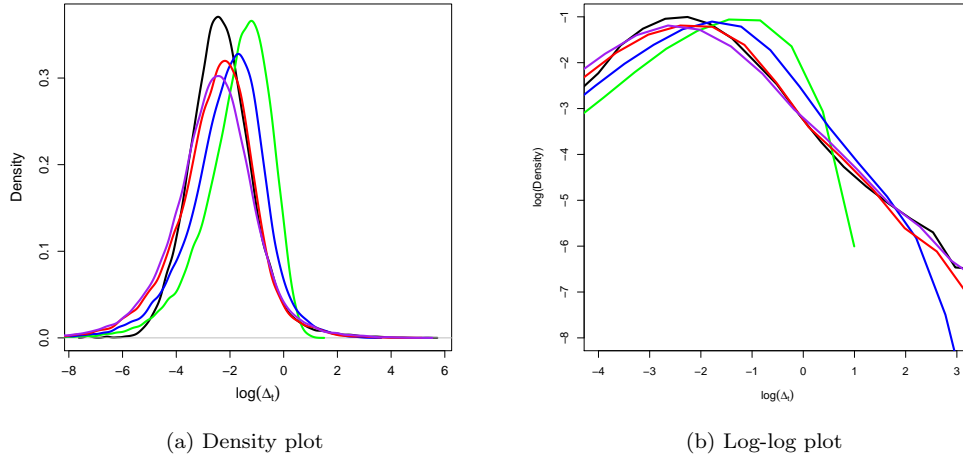


FIG. 6: The black line shows the empirical waiting time distribution $p(\log(\Delta))$ from the solar flare catalog. The colored lines show the waiting time distributions that correspond to the fitted homogenous Poisson process (green), change point Poisson process (red), seasonal Poisson process (blue), and Lognormal Hawkes process (purple).

To avoid over-interpretation of our findings, we do not claim that the above Hawkes process provides a complete and correct model for solar flare occurrences. There will doubtless be many second-order corrections which could be added into the relatively simple specification from Equation 10. Indeed, Figure 6 shows how the waiting time distribution under both the Hawkes process and the change point model compares to the empirically observed waiting times from the GOES catalog. It can be seen that while both seem to capture the power law tails, they do seem to over-predict the shorter waiting times. The fact that the change point specification over-predicts has previously been noticed by [3] and others, who speculated that it may be a result of shorter waiting times being missed in the GOES cata due to potential missing data caused by flares overlapping in time. Nonetheless, the fact that such a simple model with a small number of parameters is able to capture most of the features of solar flare occurrences, which were previously believed to require enormous numbers of parameters, is significant and in our opinion points towards a substantial non-Poissonian element. From the same plot, it can also be seen that both the homogenous Poisson and the seasonal Poisson models complete fail to capture the empirical distribution, with the seasonal model failing to reproduce the power-law tails altogether.

VI. CONCLUSION

Understanding the solar flare waiting time distribution is an important first step in forecasting the occurrence times of large flares. It also links into questions concerning the underlying physical model of solar flare production, since physical theories such as the avalanche model imply particular types of distributions. In this work we have argued that the commonly used Poissonian model for solar flares does not seem to capture the observed features of the empirical waiting time distribution.

The periodic nature of the solar cycle means that a non-stationary Poisson process must be used, such as one which has sinusoidal seasonal components, or the more commonly used change point model. But despite having a large number of free parameters, these models seem to give a systematically worse fit to the observed data than a simple self-exciting model which discards the Poissonity assumption and allows for strong dependence between the occurrence of successive flares. Despite having only a small number of parameters, this self-exciting model provides an excellent fit to the data and seems to capture the observed characteristics of the empirical waiting time distribution. Since similar models are routinely used for describing and forecasting the occurrence of other physical processes such as earthquakes and tsunamis, this also points towards an interesting universality component for this class of model.

[1] J. P. Eastwood, E. Biffis, M. A. Hapgood, L. Green, M. M. Bisi, R. D. Bentley, R. Wicks, L.-A. McKinnell, M. Gibbs, and C. Burnett, *Risk Analysis* **37**, 206 (2017).

[2] D. N. Baker, E. Daly, I. Daglis, J. G. Kappenman, and M. Panasyuk, *Space Weather* **2** (2004), 10.1029/2003SW000044.

- [3] M. S. Wheatland, *The Astrophysical Journal* **536**, L109 (2000).
- [4] C. Li, S. J. Zhong, Z. G. Xu, H. He, Y. Yan, P. F. Chen, and C. Fang, *Monthly Notices of the Royal Astronomical Society: Letters* **479**, L139 (2018).
- [5] F. Lepreti, V. Carbone, and P. Veltri, *The Astrophysical Journal Letters* **555**, L133 (2001).
- [6] A. Greco, W. H. Matthaeus, S. Servidio, and P. Dmitruk, *Physical Review E* **80**, 046401 (2009).
- [7] D. Telloni, V. Carbone, F. Lepreti, and E. Antonucci, *AIP Conference Proceedings* **1720**, 020005 (2016).
- [8] M. Wheatland, *Solar Physics* **214**, 361 (2003).
- [9] M. S. Wheatland, *Space Weather* **3**, n/a (2005).
- [10] S. Elvidge and M. J. Angling, *Space Weather* **16**, 417 (2018).
- [11] A. Anastasiadis, A. Papaioannou, I. Sandberg, M. Georgoulis, K. Tziotziou, A. Kouloumvakos, and P. Jiggins, *Solar Physics* **292** (2017).
- [12] V. De la Luz, E. P. Balanzario, and T. Tsiftsi, *Solar Physics* **293** (2018).
- [13] M. J. Aschwanden and J. M. McTiernan, *The Astrophysical Journal* **717**, 683 (2010).
- [14] C. Li, S. J. Zhong, L. Wang, W. Su, and C. Fang, *The Astrophysical Journal Letters* **792**, L26 (2014).
- [15] E. L. Geist, *Bulletin of the Seismological Society of America* **104**, 13 (2014), iP-052031.
- [16] S. Xiao, A. Kottas, and B. Sans, *The Annals of Applied Statistics* **9**, 353 (2015).
- [17] V. Filimonov and D. Sornette, *Physical Review E* **85**, 056108 (2012).
- [18] T. Omi, Y. Hirata, and K. Aihara, *Physical Review E* **96**, 012303 (2017).
- [19] C. Blundell, J. Beck, and K. A. Heller, in *Advances in Neural Information Processing Systems 25*, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc., 2012) pp. 2600–2608.
- [20] Y. Ogata, *Journal of The American Statistical Association - J AMER STATIST ASSN* **83**, 9 (1988).
- [21] D. Marsan and O. Lenglin, *Science (New York, N.Y.)* **319**, 1076 (2008).
- [22] L. de Arcangelis, C. Godano, E. Lippiello, and M. Nicodemi, *Physical Review Letters* **96**, 051102 (2006).
- [23] E. Friis-Christensen and K. Lassen, *Science* **254**, 698 (1991).
- [24] A. Clauset, C. Shalizi, and M. Newman, *SIAM Review* **51**, 661 (2009).
- [25] T. J. DiCiccio and B. Efron, *Statistical Science* **11**, 189 (1996).
- [26] D. Daley and D. Vere-Jones, **Vol. 1** (2003), 10.1007/b97277.
- [27] G. Schwarz, *The Annals of Statistics* **6**, 461 (1978).
- [28] A. R. Liddle, *Monthly Notices of the Royal Astronomical Society: Letters* **377**, L74 (2007).
- [29] D. Posada and T. R. Buckley, *Systematic Biology* **53**, 793 (2004).
- [30] R. Killick, P. Fearnhead, and I. A. Eckley, *Journal of the American Statistical Association* **107**, 1590 (2012).
- [31] A. G. Hawkes and D. Oakes, *Journal of Applied Probability* **11**, 493 (1974).
- [32] E. Brown, R. Barbiere, V. Ventura, and L. Frank, *Neural Computation* **14**, 325 (2001).